Bias of Estimates of Secondary Parameters in Linear-Boundary Sequential Tests

BY VLADIMIR DRAGALIN AND BENJAMIN YAKIR AND W.J. HALL

GlaxoSmithKline and the Hebrew University and the University of Rochester

The context is that of a sequential trial based on Brownian motion with linear stopping boundaries, possibly truncated. Along with the monitoring process, a secondary Gaussian process with constant mean is observed; the mean is to be estimated once the monitoring process reaches a boundary. We provide a formula for the conditional bias, conditioning on the final position of the monitoring process; this formula can then be integrated to obtain an overall bias. Special attention is given to evaluating bias – mathematically and by Monte Carlo – of the Kaplan-Meier estimator of one of the survival functions (and similarly for the Nelson-Aalen estimator of the corresponding cumulative hazard function) upon completion of a survival-analysis-based two-arm clinical trial. Implications in a recent clinical trial are cited.

1. Introduction. It is now standard practice in large clinical trials to have interim monitoring and the potential for early stopping—*sequential clinical trials.* Statistical issues then arise in the interpretation of the data, issues connected with the bias that inevitably enters in statistical procedures that were designed for fixed-sample (*nonsequential*) trials when used to interpret data from sequential trials. Substantial advances have been made in recent years to eliminate such bias in the primary inference: deciding whether or not there is a statistically significant effect (with significance quantified by an appropriately adjusted p-value). Progress has also been made in removing bias from estimators of the primary parameter, but some issues remain. Relatively little progress has been made in evaluating, or removing, bias in inference about other parameters, whether testing hypotheses about them or quantifying their possible magnitude. Examples of such secondary inference are

- Estimating the cumulative hazard function in a specific arm, using the Nelson-Aalen statistic
- Estimating the survival function in a specific arm, using the Kaplan-Meier statistic
- Estimating area between two survival curves
- Testing for a strata×treatment interaction in a trial in which recruitment is stratified
- Inference regarding regression coefficients in a Cox regression model.

2. The Problem. The monitoring process $Z(t)$ is assumed to be a Brownian motion with drift θ and unit variance per unit time. The secondary process $W(t)$ is

Gaussian, jointly with $Z(t)$, with a constant mean δ. The covariance function of W and the joint covariance function $\mathrm{cov}[Z(s), W(t)]$ complete the definition. These latter two functions are assumed to be known (parameter free); θ is the parameter of primary interest in a sequential test, while δ is a secondary parameter to be estimated when the test terminates.

The monitoring process has linear stopping boundaries:

$$\begin{aligned} &upper(U): \quad a_1 + b_1 t,\ t \le t_o; \\ &lower(L): \quad a_0 + b_0 t,\ t \le t_o; \\ &vertical(V): t = t_o \end{aligned}$$

for $a_0 < 0 < a_1, t_o \le +\infty$, and $b_1 \le b_0$ whenever $t_o = \infty$; let T be the boundary-hitting time.

For any fixed time t, $W(t)$ is an unbiased estimator (and MLE) of the secondary parameter δ. And $W(T)$ continues to be the MLE of δ when T is the hitting time of stopping boundaries, but typically a biased estimator. It follows from Theorem 2 of [9] that the bias of $W(T)$ is

$$E\Big(\int_0^T \beta(s,T)\, d[Z(s) - \theta s]\Big), \tag{1}$$

where

$$\beta(s,t) = \frac{\partial}{\partial s}\, \mathrm{cov}[Z(s), W(t)].$$

We also define $T_U (\le \infty)$ as the first time the monitoring process hits the upper linear boundary (ignoring other boundaries), and T_L similarly for the lower linear boundary. Write $B(s) = Z(s) - \theta s$, a standard Brownian motion. In terms of $B(s)$, stopping times are as above but with the upper and lower boundary slopes shifted, from b_i to $b_i - \theta$ $(i = 0, 1)$.

We propose computing the bias (1) by first conditioning on the coordinates of the monitoring process when first hitting the stopping boundary, and later integrating with respect to the distribution along the stopping boundary. Expressions for the density along linear boundaries are available [1, 3]. Thus, in principle, the only problem is to compute the conditional expectation of $\int_0^T \beta(s,T)\, dB(s)$, given $(T, B(T))$ with T the hitting time of the shifted boundaries for Brownian motion B. The only random element here is the process $B(s)$, with the conditional measure. (If conditioning only on the first position, the measure would be Brownian bridge, but conditioning also implies that path lies between stopping boundaries.)

3. The Conditional Bias. For general linear boundaries, the methods below can be developed to provide a Volterra integral equation of the second kind for the conditional bias. We consider first a simpler case where only one boundary (say, the upper) is present, and indicate extensions and approximations for the general case afterwards. Then the boundary location (t, x) satisfies $x = a_1 + (b_1 - \theta)t$, say

$a + bt$, and hence it is sufficient to condition on $T\ (= T_U) = t$. We apply results and methods from [9] to compute the desired conditional expectation

$$\nu_U(t) = E\Big(\int_0^t \beta(s,t)\, dB(s) \;\Big|\; T_U = t \Big).$$

Define

$$M_\lambda(t) = E\Big(\exp\{i\lambda \int_0^t \beta(s,t)\, dB(s)\} \;\Big|\; T_U = t \Big).$$

Note that

$$\frac{1}{i}\frac{d}{d\lambda} M_\lambda(t)\Big|_{\lambda=0} = \nu_U(t).$$

According to Theorem 3 of [9], the characteristic function $M_\lambda(t)$ satisfies

$$M_\lambda(t) = \frac{\gamma(t)K_\lambda(t) - \int_0^t M_\lambda(s)\gamma(s,t)K_\lambda(s,t)f_a(s,t)\, ds}{c(t) - \int_0^t c(s,t)f_a(s,t)\, ds}, \tag{2}$$

where, for $s < t$ (and some changes in notation),

$$\bar\beta(s,t) = \frac{1}{t-s}\int_s^t \beta(\tau,t)\, d\tau,$$

$$\sigma_\beta(s,t) = \int_s^t [\beta(\tau,t) - \bar\beta(s,t)]^2\, d\tau,$$

$$K_\lambda(t) = \exp\Big\{ i\lambda(a+bt)\bar\beta(0,t) - \frac{1}{2}\lambda^2\sigma_\beta(0,t)\Big\},$$

$$K_\lambda(s,t) = \exp\Big\{ i\lambda b(t-s)\bar\beta(s,t) - \frac{1}{2}\lambda^2\sigma_\beta(s,t)\Big\},$$

$$c(t) = (a+bt)/t = (a/t) + b,$$

$$c(s,t) = \frac{(a+bt) - (a+bs)}{t-s} = b,$$

$$\gamma(s,t) = c(s,t) - i\lambda\bar\beta(s,t),$$

$$\gamma(t) = c(t) - i\lambda\bar\beta(0,t),$$

and where $f_a(s,t)$ is the conditional density of T_U at $s\ (< t)$, given $B(t) = a + bt$. From Brownian bridge considerations, we find

$$f_a(s,t) = \frac{a}{s^2}\sqrt{\frac{st}{t-s}}\,\phi\Big(a\sqrt{\frac{t-s}{st}}\Big) = \frac{a}{s} f_{B(s)|B(t)}(a+bs|a+bt) \tag{3}$$

with ϕ the standard normal density and $f_{B(s)|B(t)}(x|y)$ the conditional (given $B(t) = y$) density of $B(s)$ at point x. It follows that the denominator in (2) is a/t. Moreover

$$\begin{aligned} \bar\beta(s,t) &= \frac{1}{t-s}\int_s^t \frac{\partial}{\partial\tau}\,\mathrm{cov}[Z(\tau), W(t)]\, d\tau \\ &= \frac{1}{t-s}\{\,\mathrm{cov}[Z(t), W(t)] - \,\mathrm{cov}[Z(s), W(t)]\}; \end{aligned} \tag{4}$$

and we write $\bar{\beta}(t) = \bar{\beta}(0,t) = \operatorname{cov}[Z(t), W(t)]/t$.

Using the relation

$$K_\lambda(t) = \int_0^t M_\lambda(s) K_\lambda(s,t) f_a(s,t)\, ds,$$

it follows from (2) that

$$M_\lambda(t) = K_\lambda(t)\Big(1 - \frac{i\lambda t}{a}\bar{\beta}(t)\Big) + \frac{i\lambda t}{a}\int_0^t \bar{\beta}(s,t) M_\lambda(s) K_\lambda(s,t) f_a(s,t)\, ds.$$

Taking derivatives with respect to λ and letting $\lambda \to 0$ gives

$$\nu_U(t) = (a+bt)\bar{\beta}(t) - \frac{t}{a}\bar{\beta}(t) + \frac{t}{a}\int_0^t \bar{\beta}(s,t) f_a(s,t)\, ds.$$

In terms of the original monitoring process and stopping boundaries, and using (4), we conclude:

PROPOSITION 1. *Let T_U be the hitting time by Z of the (single) linear boundary $a_1 + b_1 t$ $(a_1 > 0)$. Then the conditional bias of the estimator $W(T_U)$ of δ, on the event $\{T_U < \infty\}$, is*

$$\begin{aligned} \nu_U(t) &= E\big(W(T_U) | T_U = t\big) - \delta \\ &= \Big[\frac{Z(t)}{t} - \theta\Big] \operatorname{cov}[Z(t), W(t)] \\ &\quad + \int_0^t \frac{1}{t-s}\Big(s \operatorname{cov}[Z(t), W(t)] - t \operatorname{cov}[Z(s), W(t)]\Big)\frac{1}{a_1} f_{a_1}(s,t)\, ds. \end{aligned} \tag{5}$$

Similarly, with T_L the hitting time by Z of the (single) boundary $a_0 + b_0 t$ $(a_0 < 0)$, the conditional bias $\nu_L(t)$ of $W(T_L)$ on $\{T_L < \infty\}$ is given by (5) with a_1 replaced by $-a_0$.

Note that the only dependence on the slope b_1 in (5) is the implicit condition that $Z(t) = a_1 + b_1 t$. We speculate that, for many purposes, the leading term in (5) will provide a satisfactory approximation; see later sections. Indeed, the second term may vanish:

COROLLARY 1. *If, for some function $v(t)$,*

$$\operatorname{cov}[Z(s), W(t)] = \frac{s}{t} v(t) \qquad \text{for } 0 \le s \le t, \tag{6}$$

then

$$\nu_U(t) = \Big[\frac{Z(t)}{t} - \theta\Big] v(t). \tag{7}$$

If there is a vertical (truncation) boundary, at t_o, (5) remains valid on the event $\{T_U < t_o\}$; we deal with the possibility $\{T_U = t_o\}$ below.

We turn now to the general linear boundary case, but only on the event $\{T < t_o\}$—and more specifically on $\{T = T_U < t_o\}$. Write

$$p(t) = P(T_L > t | T_U = t) = p^U(t)/p_U(t),$$

the ratio of the sub-density of T on the event $\{T = T_U\}$ divided by the density of T_U. The former is given in [3] while the latter is well-known to be

$$p_U(t) = at^{-3/2}\phi(at^{-1/2} + bt^{1/2}) = \frac{a}{t}\phi_t(a + bt),$$

with ϕ_c the normal density with mean zero and variance c.

Then

$$\begin{aligned} E\Big(\int_0^t \beta\, dB \;\Big|\; T = T_U = t\Big) &= E\Big(\int_0^t \beta\, dB \;\Big|\; T_L > t, T_U = t\Big) \\ &= E\Big(\int_0^t \beta\, dB; T_L > t \;\Big|\; T_U = t\Big)\Big/p(t) \\ &= \Big\{E\Big(\int_0^t \beta\, dB \;\Big|\; T_U = t\Big) \\ &\qquad - E\Big(\int_0^t \beta\, dB; T_L < t \;\Big|\; T_U = t\Big)\Big\}\Big/p(t). \end{aligned} \tag{8}$$

Now writing $q_L(s|t)$ for the conditional density of T_L at s, given $T_U = t$,

$$E\Big(\int_0^t \beta\, dB; T_L < t \;\Big|\; T_U = t\Big) = \int_0^t E\Big(\int_0^t \beta\, dB \;\Big|\; T_L = s, T_U = t\Big) q_L(s|t)\, ds. \tag{9}$$

Integrating first over $(0, s)$ and then over (s, t),

$$\begin{aligned} E\Big(\int_0^t \beta\, dB \;\Big|\; T_L = s, T_U = t\Big) &= E\Big(\int_0^s \beta\, dB \;\Big|\; T = T_L = s\Big) \\ &\quad + E\Big(\int_s^t \beta\, dB \;\Big|\; B(s) = a_0 + (b_0 - \theta)s, T_U = t\Big). \end{aligned} \tag{10}$$

Finally, the conditional density $q_L(s|t)$ in (9) may be written as the joint density of (T_L, T_U) at (s, t) divided by the density of T_U at t; for $s < t$, the numerator is also the joint density of (T, T_U) at (s, t), which equals $p^L(s)p_{U'(s)}(t-s)/p_U(t)$—the numerator here being the density of T at s on $\{T = T_L\}$ times the density of $T'_{U'(s)}$ at $(t - s)$ where $T'_{U'(s)}$ is the hitting time of the single boundary $a'_1 + (b_1 - \theta)u$ by a standard Brownian motion $B'(u)$ with $a'_1 = a_1 - a_0 + (b_1 - b_0)s > 0$, i.e.

$$q_L(s|t) = \frac{p^L(s)p_{U'(s)}(t-s)}{p_U(t)}.$$

Hence, (8–10) yield

$$\begin{aligned} p^U(t)E\Big(\int_0^t \beta\, dB \;\Big|\; T = T_U = t\Big) &= p_U(t)E\Big(\int_0^t \beta\, dB \;\Big|\; T_U = t\Big) \\ &\quad - \int_0^t E\Big(\int_s^t \beta\, dB \;\Big|\; B(s) = a_0 + (b_0 - \theta)s, T_U = t\Big) p_{U'(s)}(t-s)p^L(s)\, ds \end{aligned}$$

$$- \int_0^t p^L(s) E\Big(\int_0^s \beta \, dB \;\Big|\; T = T_L = s \Big) p_{U'(s)}(t-s) \, ds.$$

The conditional expectation on the left as well as the one in the last term on the right, involves two-boundary conditions, whereas the first two conditional expectations on the right involve only one-boundary condition and are given in Proposition 1.

For practical usage, we suggest substituting a one-boundary version for the last conditional expectation on the right, thereby yielding an explicit approximation for the conditional bias on the upper boundary.

On the other hand, by symmetry, a similar expression can be obtained for the conditional on $\{T = T_L = t\}$ bias of $W(T)$. Substituting this in the last integral on the right, gives a second kind Volterra integral equation for

$$V^U(t) = p^U(t) E\Big(\int_0^t \beta \, dB \;\Big|\; T = T_U = t \Big)$$

which could be solved numerically.

Let

$a(u) = a_1 - a_0 + (b_1 - b_0)u, \quad$ for $\quad u > 0;$

$p_L(s; a)$ and $p_U(s; a)$ be the densities of T_L and T_U respectively, but with the changed intercept a (instead of a_0 and a_1 respectively);

$\nu_U(t; a)$ be the $\nu_U(t)$ defined by the right hand side of (5) with the changed intercept a.

PROPOSITION 2. *The conditional bias of the estimator $W(T)$ of δ, given $\{T = T_U = t\}$, on the event $\{T < t_o\}$, is*

$$\nu^U(t) = E\Big(W(T_U) \;\Big|\; T = T_U = t \Big) - \delta = V^U(t)/p^U(t),$$

where $V^U(t)$ is the solution of the following second kind Volterra integral equation:

$$V^U(t) = f(t) + \int_0^t K(s,t) V^U(s) \, ds \tag{11}$$

with

$$\begin{aligned} f(t) = p_U(t)\nu_U(t; a_1) &- \int_0^t p_U(t-s; a(s))\nu_U(t-s; a(s))p^L(s) \, ds \\ &- \int_0^t p_L(s)\nu_U(s; -a_0)p_U(t-s; a(s))p^L(s) \, ds \\ &+ \int_0^t \Big\{ \int_0^s \nu_U(s-u; -a(u))p_L(s-u; -a(u))p^U(u) \, du \Big\} \\ &\qquad \times p_U(t-s; a(s))p^L(s) \, ds \end{aligned}$$

and

$$K(s,t) = -\int_s^t p_L(u-s; -a(s)) p_U(t-u; a(u)) p^L(u)\, du.$$

A similar result holds for the conditional bias $\nu^L(t)$ *of* $W(T)$ *given* $\{T = T_L = t\}$.

We now return to the case of boundaries consisting of upper and vertical components, but no lower boundary—a truncated one-sided SPRT. As noted above, what is yet needed is a formula for the conditional bias given that the process has not crossed the upper boundary by time t_o and is at level x $(< a + bt_o)$ at time t_o. Writing $r(t,x) = P(T_U > t \mid B(t) = x)$, we find

$$(12)\quad E\Big(\int_0^t \beta\, dB \;\Big|\; T_U > t, B(t) = x\Big) = E\Big(\int_0^t \beta\, dB; T_U > t \;\Big|\; B(t) = x\Big)/r(t,x).$$

The expectation term in the last expression in (12) is of the form dealt with in (5). Similar arguments would lead to a recursive formula for two-boundaries with a truncation—that is, general boundaries. Details are omitted.

4. A Random Sampling Model. When a Brownian motion monitoring process serves as an approximation to a simple random sampling model, as considered by Whitehead [7], a great deal of simplification occurs. Thus, we assume, that to an adequate approximation, the stopping rule is based on a cumulative sum with independent identically distributed increments reaching a linear boundary. Each increment has expectation θ. Information is therefore proportional to sample size n. A secondary process is based on another cumulative sum, with increments having expectation δ, and $W(t)$ represents the corresponding average. The correlation between increments in the two processes is a constant r. Then the covariance between the two processes satisfies the condition in the Corollary with a constant $v(t)$.

Then, as apparent in [7], and noted in [4, 5], the conditional bias of $W(t)$ is given by Corollary 1 (that is, the leading term in (5)), with a constant covariance term. Thus, the second term in (5) may be interpreted as a departure-from-random-sampling effect, while the first term relates the bias to that of $Z(T)/T$ as an estimator of the primary parameter θ. This latter fact has been used in [7, 8, 4] to derive bias adjustments, or fully unbiased estimators, of secondary parameters in a random sampling model. Whenever the second term is small, the first term in (5) may provide the basis for some bias removal from $W(T)$.

5. The Bias of a Kaplan-Meier estimator after a Survival-Analysis-Based Sequential Trial. A random sampling model, though useful in some applications, cannot cover many situations that can occur in secondary inference. In this section, we consider estimating the survival function in a specific arm using the Kaplan-Meier estimator. We will assume that the monitoring process is the logrank statistic, plotted against its (estimated) variance. The model is such that the observations (E_{ki}, X_{ki}, U_{ki}), $k = 1, 2$, $i = 1, 2, \ldots$,—where E_{ki} is the entry time, X_{ki} is the survival time and U_{ki} is the censoring time for the i-th subject from arm

k—are independent random vectors with nonnegative components with (E_{ki}, U_{ki}) and X_{ki} independent for all i and k. Define

$$N_{ki}(t,a) = \mathbb{1}\{E_{ki} + X_{ki} \le t, X_{ki} \le U_{ki}, X_{ki} \le a\},$$

$$Y_{ki}(t,a) = \mathbb{1}\{t - E_{ki} \ge a, X_{ki} \ge a, U_{ki} \ge a\},$$

and let

$$N_k = \sum_{i=1}^{n_k} N_{ki} \quad \text{and} \quad Y_k = \sum_{i=1}^{n_k} Y_{ki}$$

be the counting process and the risk process for arm k. Let $n = n_1 + n_2$, $N_\cdot = N_1 + N_2$ and $Y_\cdot = Y_1 + Y_2$.

Assume that the limits

$$y_k(t,a) = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} P(t - E_{ki} \ge a, U_{ki} \ge a) \tag{13}$$

exist for all k, t, a and that they are continuous in t and a and positive. Assume that $n_1/n \to \gamma$ and that $0 < \gamma < 1$. Denote $y_\cdot(t,a) = \gamma y_1(t,a) + (1-\gamma) y_2(t,a)$. Let S_k, Λ_k and λ_k be the survival function, the cumulative hazard function and the hazard function respectively of the random variable X_{ki}.

A natural nonparametric estimator of the value of the cumulative hazard at survival time a, given the data that was accumulated by calendar time t, is the Nelson-Aalen estimator

$$\hat{\Lambda}_k^{(n)}(t,a) = \int_0^a \frac{J_k(t,s)}{Y_k(t,s)} N_k(t, ds)$$

where $J_k(t,s) = \mathbb{1}\{Y_k(t,s) > 0\}$. The Kaplan-Meier estimator is given by

$$\hat{S}_k^{(n)}(t,a) = \prod_{s\le a} [1 - \Delta\hat{\Lambda}_k^{(n)}(t,s)] = \prod_{s\le a} \Big(1 - \frac{J_k(t,s)}{Y_k(t,s)} \Delta N_k(t,s)\Big) \tag{14}$$

where

$$\Delta\hat{\Lambda}_k^{(n)}(t,s) = \hat{\Lambda}_k^{(n)}(t,s) - \hat{\Lambda}_k^{(n)}(t,s_-)],$$

and similarly for ΔN_k.

The logrank statistic can be defined as

$$\hat{Z}^{(n)}(t) = \int_0^t \frac{Y_1^{(n)}(t,s) Y_2^{(n)}(t,s)}{Y_\cdot^{(n)}(t,s)} \Big[\hat{\Lambda}_1^{(n)}(t, ds) - \hat{\Lambda}_2^{(n)}(t, ds)\Big]. \tag{15}$$

Gu and Lai [2] proved the following:

For fixed $S_1 = S$ (and therefore $\lambda_1 = \lambda$), suppose that as $n \to \infty$, $S_2 \to S$ such that

$$\int_0^\tau |\lambda_2(s)/\lambda(s) - 1| \lambda(s)\, ds = O(n^{-1/2})$$

and

$$\sqrt{n}\{(\lambda_2(s)/\lambda(s)) - 1\} \to g(s),$$

uniformly in $s \in I$ *and* $\sup_{s\in I} |g(s)| < \infty$ *for all closed intervals* I *of* $\{s \in [0,\tau] : S(s) > 0\}$. *Then* $\{n^{-1/2}\hat{Z}^{(n)}(t), 0 \le t \le \tau\}$ *converges weakly in* $D[0,\tau]$ *to* $\{Z(t), 0 \le t \le \tau\}$, *where* $Z(t)$ *is a Gaussian process with independent increments with drift*

$$\mu(t) = \gamma(1-\gamma)\int_0^t \frac{y_1(t,u)y_2(t,u)}{y_\cdot(t,u)} g(u)\, dS(u)$$

and variance

$$Var(Z(t)) = -\gamma(1-\gamma)\int_0^t \frac{y_1(t,u)y_2(t,u)}{y_\cdot(t,u)}\, dS(u).$$

Under the Cox proportional hazards model with contiguous alternatives, i.e. $g(u) \equiv -\theta$, the limiting process Z is a time-transformed Brownian motion with drift θ and unit variance per unit time.

The (asymptotic) covariance structure between the monitoring process and the Nelson-Aalen estimator for the group specific a-years cumulative hazard is given by:

$$\begin{aligned}\operatorname{cov}[Z(s), \Lambda_1(t,a)] &= \int_0^{s\wedge a} \frac{y_2(s,u)}{y_\cdot(s,u)} \frac{y_1(s,u)}{y_1(t,u)} \lambda(u)\, du \\ &= (1-\gamma)\int_0^{s\wedge a} \frac{G(s,u)}{G(t,u)} \lambda(u)\, du, \\ \operatorname{cov}[Z(s), \Lambda_2(t,a)] &= -\gamma \int_0^{s\wedge a} \frac{G(s,u)}{G(t,u)} \lambda(u)\, du\end{aligned}$$

for $s \le t$, $t > a$. Here we assume $G(t,u) = P(E_1 \le t-u, U_1 \ge u) = P(E_2 \le t-u, U_2 \ge u)$.

The random field $n^{1/2}[\hat{S}_k^{(n)}(t,a) - S_k(a)]$ is asymptotically equivalent to the random field $-S_k(a)n^{1/2}(\hat{\Lambda}_k^{(n)}(t,a) - \Lambda_k(a))$. It follows that

$$\begin{aligned}\operatorname{cov}[Z(s), S_1(t,a)] &= -(1-\gamma)S_1(a)\int_0^{s\wedge a} \frac{G(s,u)}{G(t,u)} \lambda(u)\, du, \\ \operatorname{cov}[Z(s), S_2(t,a)] &= \gamma S_2(a)\int_0^{s\wedge a} \frac{G(s,u)}{G(t,u)} \lambda(u)\, du.\end{aligned}$$

This suggests the following adjustments to the Kaplan-Meier estimators when the trial was stopped at $T = t$:

$$\tilde{S}_1(t,a) = \hat{S}_1(t,a) - (1-\gamma)\Big(\frac{\hat{Z}(t)}{t} - \tilde{\theta}\Big)\hat{S}_1(t,a)\log \hat{S}_1(t,a) \tag{16}$$

and

$$\tilde{S}_2(t,a) = \hat{S}_2(t,a) + \gamma\Big(\frac{\hat{Z}(t)}{t} - \tilde{\theta}\Big)\hat{S}_2(t,a)\log \hat{S}_2(t,a), \tag{17}$$

where $\bar{\theta}$ is a "good" estimator of the primary parameter θ, e.g., uniformly minimum variance mean-unbiased estimator (UMVUE), or median-unbiased estimator, or bias adjusted MLE.

6. Simulation Results. In order to evaluate the accuracy of these adjustments to Kaplan-Meier estimators, we performed Monte Carlo studies. First, we demonstrate that the original (naive, ignoring the sequential stopping rule) Kaplan-Meier estimators are indeed biased. Then we show that the adjustments proposed in (16) and (17), although based on the crude approximation of conditional bias of the two-boundary stopping rule with only the first term of the conditional bias of a one-boundary stopping rule (5), are accurate enough.

But before that, we need to introduce an appropriate measure of bias of these estimators. One can use *overall bias* averaging the estimators *over all* simulated trials. It can be argued that the conditional distribution, given that the trial lasted t_0 units of time, is a more appropriate measure under which the properties of an estimator should be analyzed, e.g. one would expect that the estimators calculated after early-stopped trials will perform differently from the estimators calculated after late-stopped trials. Therefore, we propose the following definition of the *conditional bias* of an estimator of survival function $S_k(a)$ at point a upon completion of the clinical trial:

$$E\Big(\hat{S}_k(T,a) \;\Big|\; T \geq t_0\Big) - S_k(a),$$

for given t_0. A "good" estimator would have small both overall and conditional biases over a reasonable range of t_0 values.

To assess the bias of the naive Kaplan-Meier estimator and the proposed adjusted one under each of these measures, data arising from a randomized controlled clinical trial of a new treatment are generated. Upon entering the study, patients are randomly assigned to one of the two groups, receiving either the new treatment (TRT arm) or the control treatment (CT arm). Assume that patients enter the trial uniformly during the two-year recruitment period. The primary endpoint of the clinical trial is patient survival. The monitoring process is the logrank statistic based on observations from the two arms. The proportional hazard model is assumed. A triangular stopping boundary test is used for testing the null hypothesis of no difference between treatments ($\theta = 0$, where θ is the log hazard ratio) versus the one-sided alternative that the new treatment is better with a 5% significance level and 95% power at $\theta_1 = 0.63$. The interim analysis is performed after each month starting 6 months after the beginning of the trial. 1200 patients are expected to be enrolled. This is a prototype of an ongoing Multicenter Automatic Defibrillator Implantation Trial MADIT II. (By the time of revision, this trial has been successfully terminated and results reported in The New England Journal of Medicine, see [6].)

The simulations were performed for three situations: under the alternative hypothesis θ_1, under the null hypothesis $\theta = 0$ and under the in-between hypothesis $\theta_1/2$ which corresponds to the largest average duration of the trial. The survival distributions were generated from the exponential distribution with 19% two-year mortality rate in the control arm and the respective one for the treatment arm.

The average duration of the trial was 27 months under both null and alternative hypotheses and 31.5 months under the in-between hypothesis. 5000 such clinical trials were simulated under each situation.

Results are presented in Figures 1–6. The overall bias in the naive Kaplan-Meier estimator was negligible under the third situation. Under the alternative hypothesis, the naive Kaplan-Meier estimator was slightly underestimating the control arm survival at the early time points and overestimating it (even considerably) at the late time points. The situation was reversed for the treatment arm. Similar behavior (but in the opposite direction) was observed under the null hypothesis. However, confining attention to late-stopped trials (those stopping after 24 months) there was considerable bias for naive Kaplan-Meier estimator over the whole range of time points. Late stopping led to over-estimating survival in the control arm and under-estimating it in the treatment arm under the alternative hypothesis, with the reverse tendency under the null hypothesis.

On the other hand, the adjustment with $\tilde{\theta} \equiv \theta$ works perfectly in all situations except under the in-between hypothesis. In this case, the stopping time is close to the vertex and thus the stopping rule effect is not so crucial, leading to an almost unbiased estimate of the primary parameter θ, which implies that the adjustment is negligible, see (16) and (17). Fortunately, the naive Kaplan-Meier estimator is working well in this situation (see Figures 5 and 6), so there is no practical need for adjustment at all.

Unfortunately, the situation is not so optimistic when a "good" estimator $\tilde{\theta}$ is used in (16) and (17). See Figure 7 and 8 (the alternative hypothesis situation). One explanation is that this estimators of θ are "good" in the ideal framework of the Brownian motion. In these simulations (like in the real trial), the monitoring was done on logrank statistic which is only approximately a Brownian motion. Furthermore the interim analysis was done after each month, not continuously. The known effect of excess over the boundary in discrete time monitoring may well apply in this situation. For example, we found, by limited simulation studies, that the actual bias of the UMVUE in this situation is of the similar magnitude as that of MLE, only in the opposite direction. We are working now on the discrete time adjustments for UMVUE and median unbiased estimators of the primary parameter. Another explanation could be the high correlation between the MLE $Z(T)/T$ and $\tilde{\theta}$, diminishing the adjustment in (16) and (17).

REFERENCES

[1] T.W. Anderson. A modification of the sequential probability ratio test to reduce the sample size. *Annals of Mathematical Statistics*, 31:165–197, 1960.

[2] M.G. Gu and T.L. Lai. Weak convergence of time-sequential rank statistics with applications to sequential testing in clinical trials. *Annals of Statistics*, 19:1403–1433, 1991.

[3] W.J. Hall. The distribution of Brownian motion on linear stopping boundaries. *Sequential Analysis*, 16:345–352, 1997. Addendum in 17: 123-124.

[4] A. Liu and W.J. Hall. Minimum variance unbiased estimation of the drift of Brownian motion with linear stopping boundaries. *Sequential Analysis*, 17:91–107, 1998.

[5] A. Liu and W.J. Hall. Unbiased estimation of secondary parameters following a sequential test. *Biometrika*, 88:895–900, 2001.

Figure 1: Bias of Kaplan-Meier estimators, $\theta = \theta_1$. Est. $\tilde{\theta} \equiv \theta$.

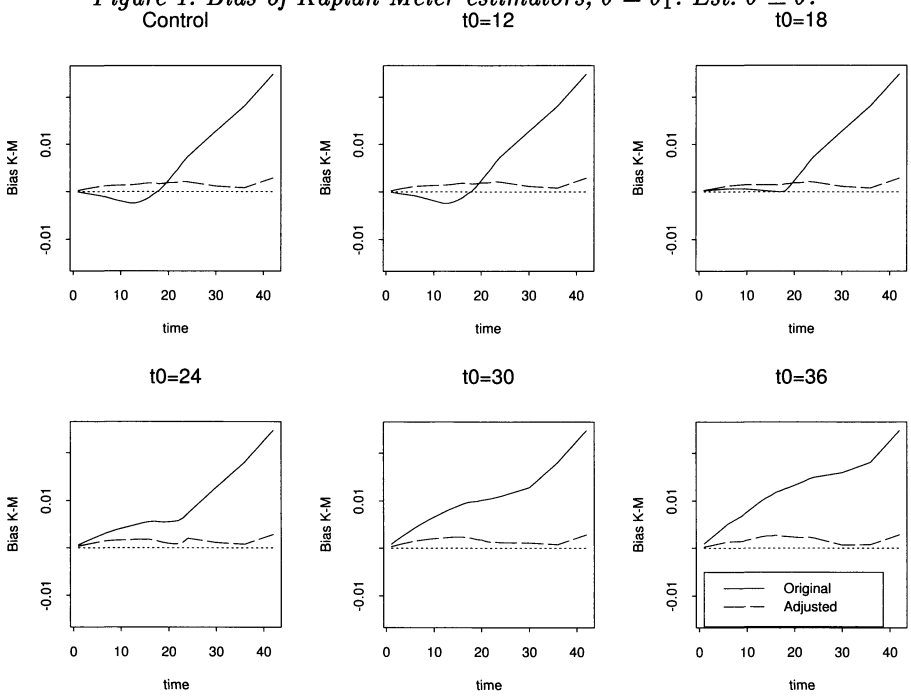


Figure 2: Bias of Kaplan-Meier estimators, $\theta = \theta_1$. Est. $\tilde{\theta} \equiv \theta$.

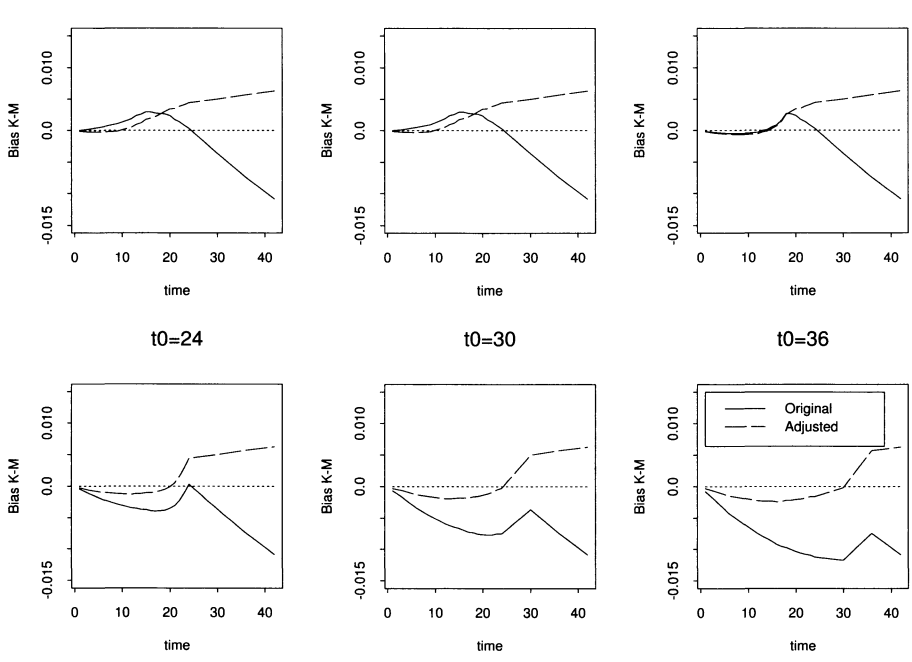


Figure 3: Bias of Kaplan-Meier estimators, $\theta = 0$. Est. $\tilde{\theta} \equiv \theta$.

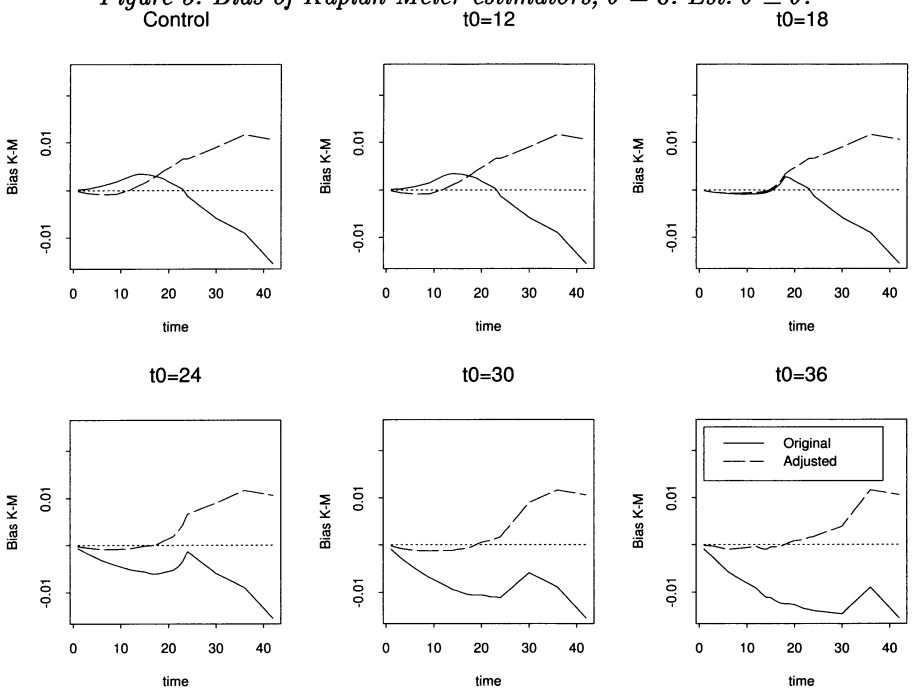


Figure 4: Bias of Kaplan-Meier estimators, $\theta = 0$. Est. $\tilde{\theta} \equiv \theta$.

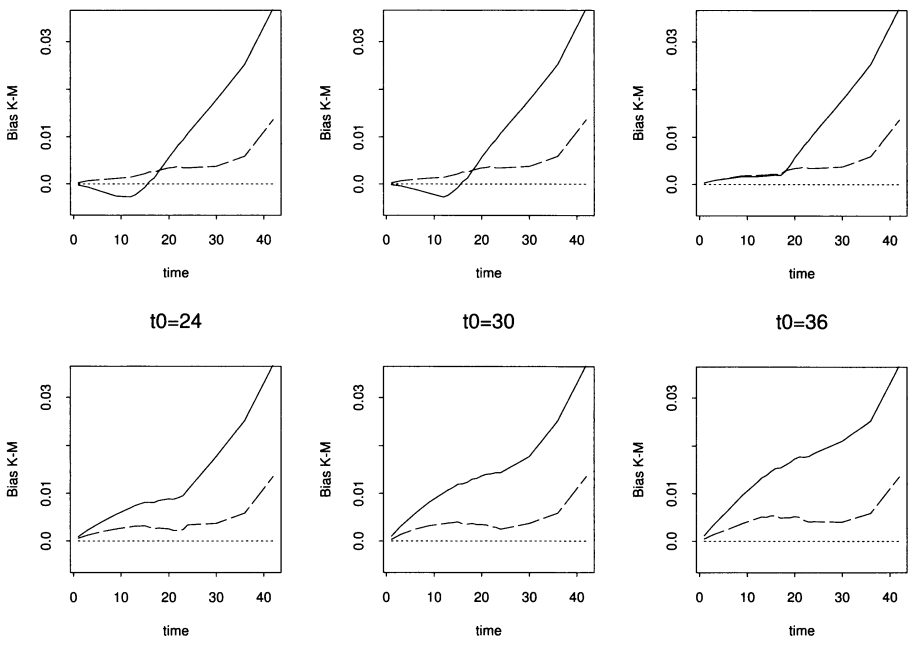


Figure 5: Bias of Kaplan-Meier estimators, $\theta = \theta_1/2$. Est. $\tilde{\theta} \equiv \theta$.

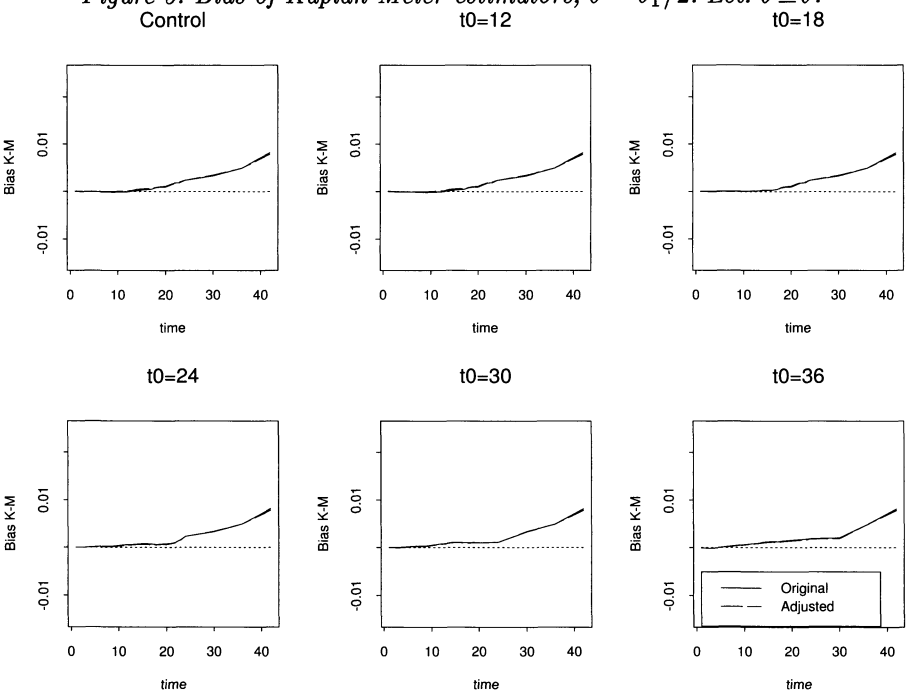


Figure 6: Bias of Kaplan-Meier estimators, $\theta = \theta_1/2$. Est. $\tilde{\theta} \equiv \theta$.

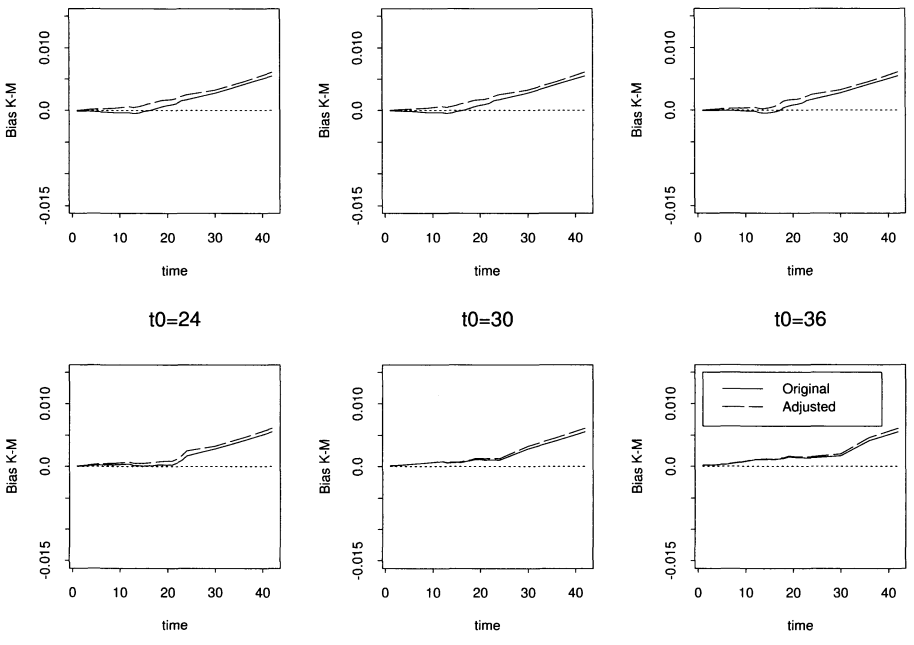


Figure 7: Bias of Kaplan-Meier estimators, $\theta = \theta_1$. Est. $\tilde{\theta}$ is UMVUE.

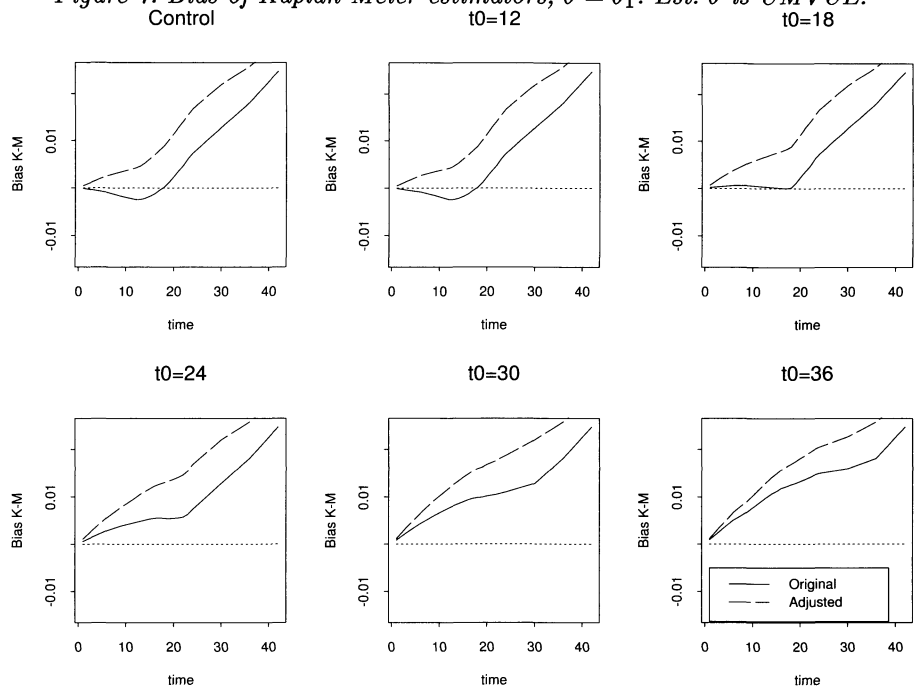
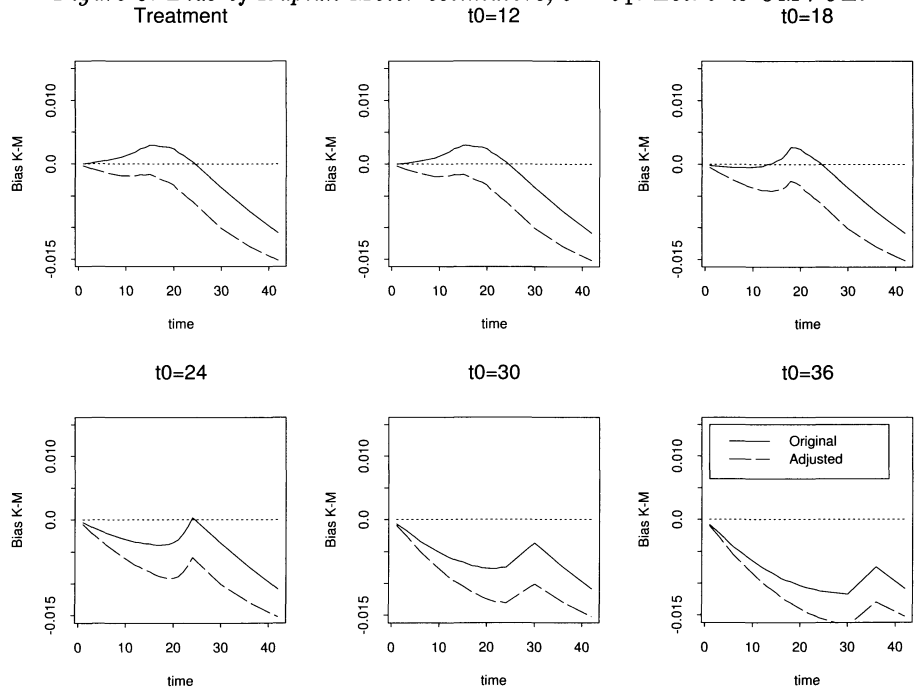


Figure 8: Bias of Kaplan-Meier estimators, $\theta = \theta_1$. Est. $\tilde{\theta}$ is UMVUE.



[6] A.J. Moss, W. Zareba, W.J. Hall, H. Klein, D.J. Wilber, D.S. Cannom, J.P. Daubert, S.L. Higgins, M.W. Brown, and M.L. Andrews. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *New England Journal of Medicine*, 346:877–883, 2002.

[7] J. Whitehead. Supplementary analysis at the conclusion of a sequential clinical trial. *Biometrika*, 42:461–471, 1986.

[8] J. Whitehead, S. Todd, and W.J. Hall. Confidence intervals for secondary parameters following a sequential test. *Journal of the Royal Statistical Society Series B*, 62:731–745, 2000.

[9] B. Yakir. On the distribution of a concomitant statistic in a sequential trial. *Sequential Analysis*, 16:287–294, 1997.

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF ROCHESTER
ROCHESTER, NY 14642-8630
hall@bst.rochester.edu

DEPARTMENT OF STATISTICS
THE HEBREW UNIVERSITY
JERUSALEM 91905, ISRAEL
msby@mscc.huji.ac.il

RESEARCH STATISTICS UNIT, BDS
GLAXOSMITHKLINE
COLLEGEVILLE, PA 19426-0989
Vladimir.2.Dragalin@gsk.com